

16C1



OICE

RECEIVED

FEB 12 2002

TECH CENTER 1600/2900

#4

1631

RAW SEQUENCE LISTING
 PATENT APPLICATION: US/09/923,444A

DATE: 02/06/2002
 TIME: 08:56:41

Input Set : N:\Crf3\RULE60\09923444A.txt
 Output Set: N:\CRF3\02062002\I923444A.raw

SEQUENCE LISTING

4 (1) GENERAL INFORMATION:
 6 (i) APPLICANT: LI, YI
 7 FLEISCHMANN, ROBERT
 9 (ii) TITLE OF INVENTION: NEUROTRANSMITTER TRANSPORTER
 11 (iii) NUMBER OF SEQUENCES: 6
 13 (iv) CORRESPONDENCE ADDRESS:
 14 (A) ADDRESSEE: Human Genome Sciences, Inc.
 15 (B) STREET: 9410 Key West Avenue
 16 (C) CITY: Rockville
 17 (D) STATE: MD
 18 (E) COUNTRY: US
 19 (F) ZIP: 20850
 21 (v) COMPUTER READABLE FORM:
 22 (A) MEDIUM TYPE: Floppy disk
 23 (B) COMPUTER: IBM PC compatible
 24 (C) OPERATING SYSTEM: PC-DOS/MS-DOS
 25 (D) SOFTWARE: PatentIn Release #1.0, Version #1.30
 27 (vi) CURRENT APPLICATION DATA:
 C--> 28 (A) APPLICATION NUMBER: US/09/923,444A
 C--> 29 (B) FILING DATE: 08-Aug-2001
 30 (C) CLASSIFICATION:
 32 (vii) PRIOR APPLICATION DATA:
 33 (A) APPLICATION NUMBER: 09/062,815
 34 (B) FILING DATE: 199-12-20
 36 (viii) ATTORNEY/AGENT INFORMATION:
 37 (A) NAME: Michele M. Wales
 38 (B) REGISTRATION NUMBER: 43,975
 39 (C) REFERENCE/DOCKET NUMBER: PF116
 41 (ix) TELECOMMUNICATION INFORMATION:
 42 (A) TELEPHONE: (301) 309-8504
 43 (B) TELEFAX: (301) 309-8439
 46 (2) INFORMATION FOR SEQ ID NO: 1:
 48 (i) SEQUENCE CHARACTERISTICS:
 49 (A) LENGTH: 2485 base pairs
 50 (B) TYPE: nucleic acid
 51 (C) STRANDEDNESS: single
 52 (D) TOPOLOGY: linear
 54 (ii) MOLECULE TYPE: DNA (genomic)
 57 (ix) FEATURE:
 58 (A) NAME/KEY: CDS
 59 (B) LOCATION: 266..2446
 62 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1:

ENTERED

RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/923,444A

DATE: 02/06/2002

TIME: 08:56:41

Input Set : N:\Crf3\RULE60\09923444A.txt

Output Set: N:\CRF3\02062002\I923444A.raw

64	CGGAGGCAGG	GAGTGAGGAG	CGAGCGGAGT	CGCGTGCGCC	GGCGCGAGCT	CCGGGTCGCC	60
66	CCAGCCCCAG	CCGGGGGCCT	GTGGCGGGGG	AGGAGCTGTG	CGTCCGCGAC	CCGTCCGGAT	120
68	CGCAGCTGCT	CGGCCGGAGT	GCACGGGCGG	AGTCTGCGCG	ACTACCCACG	CGTGACAGGT	180
70	CCCTGAATGA	GAAGGAGCTG	ACAGCAGCTG	AATTCCATCT	TCTCTGTGTG	CTGGGGAGCA	240
72	GGGCTACACG	GCCCAGGTGG	CATCA	ATG	CCG	AAG AAC AGC AAA GTG ACC CAG	292
73			Met	Pro	Lys	Asn Ser Lys Val Thr Gln	
74			1		5		
76	CGT	GAG	CAC	AGC	AGT	GAG CAT GTC ACT GAG TCC GTG GCC GAC CTG CTG	340
77	Arg	Glu	His	Ser	Ser	Glu His Val Thr Glu Ser Val Ala Asp Leu Leu	
78	10			15		20	25
80	GCC	CTC	GAG	GAG	CCT	GTG GAC TAT AAG CAG AGT GTA CTG AAT GTG GCT	388
81	Ala	Leu	Glu	Glu	Pro	Val Asp Tyr Lys Gln Ser Val Leu Asn Val Ala	
82			30		35	40	
84	GGT	GAG	GCA	GGC	GGC	AAG CAG AAG GCG GTG GAG GAG GAG CTG GAT GCA	436
85	Gly	Glu	Ala	Gly	Gly	Lys Gln Lys Ala Val Glu Glu Glu Leu Asp Ala	
86		45		50		55	
88	GAG	GAC	CGG	CCG	GCC	TGG AAC AGT AAG CTG CAG TAC ATC CTG GCC CAG	484
89	Glu	Asp	Arg	Pro	Ala	Trp Asn Ser Lys Leu Gln Tyr Ile Leu Ala Gln	
90		60		65		70	
92	ATT	GGC	TTC	TCT	GTG	GGC CTC GGC AAC ATC TGG AGG TTC CCC TAC CTG	532
93	Ile	Gly	Phe	Ser	Val	Gly Leu Gly Asn Ile Trp Arg Phe Pro Tyr Leu	
94		75		80		85	
96	TGC	CAG	AAA	AAT	GGA	GGA GGT GCT TAC CTG GTG CCC TAC CTG GTG CTG	580
97	Cys	Gln	Lys	Asn	Gly	Gly Gly Ala Tyr Leu Val Pro Tyr Leu Val Leu	
98	90		95		100	105	
100	CTG	ATC	ATC	ATC	GGG	ATC CCC CTC TTC TTC CTG GAG CTG GCT GTG GGT	628
101	Leu	Ile	Ile	Ile	Gly	Ile Pro Leu Phe Phe Leu Glu Leu Ala Val Gly	
102			110		115	120	
104	CAG	AGG	ATC	CGC	CGC	GGA AGC ATC GGT GTG TGG CAC TAT ATA TGT CCC	676
105	Gln	Arg	Ile	Arg	Arg	Gly Ser Ile Gly Val Trp His Tyr Ile Cys Pro	
106			125		130	135	
108	CGC	CTG	GGG	GGG	ATC	GGC TTC TCC AGC TGC ATA GTC TGT CTC TTT GTG	724
109	Arg	Leu	Gly	Gly	Ile	Gly Phe Ser Cys Ile Val Cys Leu Phe Val	
110		140		145		150	
112	GGG	CTG	TAT	TAT	AAT	GTG ATC ATC GGG TGG AGC ATC TTC TAT TTC TTC	772
113	Gly	Leu	Tyr	Tyr	Asn	Val Ile Ile Gly Trp Ser Ile Phe Tyr Phe Phe	
114		155		160		165	
116	AAG	TCC	TTC	CAG	TAC	CCG CTG CCC TGG AGT GAA TGT CCT GTC GTC AGG	820
117	Lys	Ser	Phe	Gln	Tyr	Pro Leu Pro Trp Ser Glu Cys Pro Val Val Arg	
118	170		175		180	185	
120	AAT	GGG	AGC	GTC	GCA	GTG GTG GAG GCA GAG TGT GAA AAG AGC TCA GCC	868
121	Asn	Gly	Ser	Val	Ala	Val Val Glu Ala Glu Cys Glu Lys Ser Ser Ala	
122			190		195	200	
124	ACT	ACC	TAC	TTC	TGG	TAC CGA GAG GCT TTG GAC ATC TCT GAC TCC ATC	916
125	Thr	Thr	Tyr	Phe	Trp	Tyr Arg Glu Ala Leu Asp Ile Ser Asp Ser Ile	
126		205		210		215	
128	TCG	GAG	AGT	GGG	GGC	CTC AAC TGG AAG ATG ACC CTG TGC CTC CTC GTG	964
129	Ser	Glu	Ser	Gly	Gly	Leu Asn Trp Lys Met Thr Leu Cys Leu Leu Val	
130		220		225		230	

RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/923,444A

DATE: 02/06/2002

TIME: 08:56:41

Input Set : N:\Crf3\RULE60\09923444A.txt

Output Set: N:\CRF3\02062002\I923444A.raw

132	GTC	TGG	AGC	ATC	GGG	GGG	ATG	GCT	GTC	GGT	AAG	GGC	ATC	CAG	TCC	TCG	1012
133	Val	Trp	Ser	Ile	Gly	Gly	Met	Ala	Val	Gly	Lys	Gly	Ile	Gln	Ser	Ser	
134		235					240					245					
136	GGG	AAG	GTG	ATG	TAT	TTC	AGC	TCC	CTC	TTC	CCC	TAC	GTG	GTG	CTG	GCC	1060
137	Gly	Lys	Val	Met	Tyr	Phe	Ser	Ser	Leu	Phe	Pro	Tyr	Val	Val	Leu	Ala	
138	250					255					260					265	
140	TGC	TTC	CTG	GTC	CGG	GGG	TTG	TTG	TTG	CGA	GGG	GCA	GTT	GAT	GGC	ATC	1108
141	Cys	Phe	Leu	Val	Arg	Gly	Leu	Leu	Leu	Arg	Gly	Ala	Val	Asp	Gly	Ile	
142					270					275					280		
144	CTA	CAC	ATG	TTC	ACT	CCC	AAG	CTG	GTC	AAG	ATG	CTG	GAC	CCC	CAG	GTG	1156
145	Leu	His	Met	Phe	Thr	Pro	Lys	Leu	Val	Lys	Met	Leu	Asp	Pro	Gln	Val	
146				285					290					295			
148	TGG	CGG	GAG	GTA	GCT	ACC	CAG	GTC	TTC	TTT	GGC	TTG	GGT	CTG	GGC	TTT	1204
149	Trp	Arg	Glu	Val	Ala	Thr	Gln	Val	Phe	Phe	Gly	Leu	Gly	Leu	Gly	Phe	
150			300					305						310			
152	GGT	GGT	GTC	ATT	GTC	TTC	TCC	AGT	TAC	AAT	AAG	CAG	GAC	AAC	AAC	TGC	1252
153	Gly	Gly	Val	Ile	Val	Phe	Ser	Ser	Tyr	Asn	Lys	Gln	Asp	Asn	Asn	Cys	
154		315					320					325					
156	CAC	TTC	GAT	GGC	GCC	CTG	GTG	TCC	TTC	ATC	AAC	TTC	TTC	ACG	TCA	GTG	1300
157	His	Phe	Asp	Gly	Ala	Leu	Val	Ser	Phe	Ile	Asn	Phe	Phe	Thr	Ser	Val	
158	330					335					340					345	
160	TTG	GCC	ACC	CTC	GTG	GTG	TTT	GTT	GTT	TTG	GGC	TTC	AAG	GCC	AAC	ATC	1348
161	Leu	Ala	Thr	Leu	Val	Val	Phe	Val	Val	Leu	Gly	Phe	Lys	Ala	Asn	Ile	
162				350						355					360		
164	ATG	AAT	GAG	AAG	TGT	GTG	GTC	GAG	AAT	GCT	GAG	AAA	ATC	CTA	GGG	TAC	1396
165	Met	Asn	Glu	Lys	Cys	Val	Val	Glu	Asn	Ala	Glu	Lys	Ile	Leu	Gly	Tyr	
166			365					370					375				
168	CTT	AAC	ACC	AAC	GTC	CTG	AGC	CGG	GAC	CTC	ATC	CCA	CCC	CAC	GTC	AAC	1444
169	Leu	Asn	Thr	Asn	Val	Leu	Ser	Arg	Asp	Leu	Ile	Pro	Pro	His	Val	Asn	
170			380					385					390				
172	TTC	TCC	CAC	CTG	ACC	ACA	AAG	GAC	TAC	ATG	GAG	ATG	GAC	AAT	GTC	ATC	1492
173	Phe	Ser	His	Leu	Thr	Thr	Lys	Asp	Tyr	Met	Glu	Met	Asp	Asn	Val	Ile	
174		395					400						405				
176	ATG	ACC	GTG	AAG	GAG	GAC	CAG	TTC	TCA	GCC	CTG	GGC	CTT	GAC	CCC	TGC	1540
177	Met	Thr	Val	Lys	Glu	Asp	Gln	Phe	Ser	Ala	Leu	Gly	Leu	Asp	Pro	Cys	
178	410				415					420					425		
180	CTT	CTG	GAG	GAC	GAG	CTG	GAC	AAG	TCC	GTG	CAG	GGC	ACA	GGC	CTG	GCC	1588
181	Leu	Leu	Glu	Asp	Glu	Leu	Asp	Lys	Ser	Val	Gln	Gly	Thr	Gly	Leu	Ala	
182				430						435					440		
184	TTC	ATC	GCC	TTC	ACT	GAG	GCC	ATG	ACG	CAC	TTC	CCC	ACC	TCC	CCG	TTC	1636
185	Phe	Ile	Ala	Phe	Thr	Glu	Ala	Met	Thr	His	Phe	Pro	Thr	Ser	Pro	Phe	
186			445					450						455			
188	TGG	TCC	GTC	ATG	TTC	TTC	TTG	ATG	CTT	ATC	AAC	CTG	GGC	CTG	GGC	AGC	1684
189	Trp	Ser	Val	Met	Phe	Phe	Leu	Met	Leu	Ile	Asn	Leu	Gly	Leu	Gly	Ser	
190			460					465					470				
192	ATG	ATC	GGG	ACC	ATG	GCA	GGC	ATC	ACC	ACG	CCC	ATC	ATC	GAC	ACC	TCC	1732
193	Met	Ile	Gly	Thr	Met	Ala	Gly	Ile	Thr	Thr	Pro	Ile	Ile	Asp	Thr	Ser	
194		475					480						485				
196	AAG	GTG	CCC	AAG	GAG	ATG	TTC	ACA	GTG	GGC	TGC	TGT	GTC	TTT	ACA	TTC	1780

RAW SEQUENCE LISTING
 PATENT APPLICATION: US/09/923,444A

DATE: 02/06/2002
 TIME: 08:56:41

Input Set : N:\Crf3\RULE60\09923444A.txt
 Output Set: N:\CRF3\02062002\I923444A.raw

```

197 Lys Val Pro Lys Glu Met Phe Thr Val Gly Cys Cys Val Phe Thr Phe
198 490 495 500 505
200 CTC GTG GGA CTG TTG TTC GTC CAG CGC TCC GGA AAC TAC TTT GTC ACC 1828
201 Leu Val Gly Leu Leu Phe Val Gln Arg Ser Gly Asn Tyr Phe Val Thr
202 510 515 520
204 ATG TTC GAT GAC TAC TCA GCC ACG CTG CCA CTC ACT CTC ATC GTC ATC 1876
205 Met Phe Asp Asp Tyr Ser Ala Thr Leu Pro Leu Thr Leu Ile Val Ile
206 525 530 535
208 CTT GAG AAC ATC GCT GTG GCC TGG ATT TAT GGA CCC AAG AAG TTC ATG 1924
209 Leu Glu Asn Ile Ala Val Ala Trp Ile Tyr Gly Pro Lys Lys Phe Met
210 540 545 550
212 CAG GAG CTG ACG GAG ATG CTG GGC TTC CGC CCC TAC CGC TTC TAT TTC 1972
213 Gln Glu Leu Thr Glu Met Leu Gly Phe Arg Pro Tyr Arg Phe Tyr Phe
214 555 560 565
216 TAC ATG TGG AAG TTC GTG TCT CCA CTA TGC ATG GCT GTG CTC ACC ACA 2020
217 Tyr Met Trp Lys Phe Val Ser Pro Leu Cys Met Ala Val Leu Thr Thr
218 570 575 580
220 GCC AGC ATC ATC CAG CTG GGG GTC ACG CCC CCG GCC TAC AGC GCC TGG 2068
221 Ala Ser Ile Ile Gln Leu Gly Val Thr Pro Pro Ala Tyr Ser Ala Trp
222 590 595 600
224 ATC AAG GAG GAG GCT GCC GAG CGC TAC CTG TAT TTC CCC AAC TGG CCC 2116
225 Ile Lys Glu Glu Ala Ala Glu Arg Tyr Leu Tyr Phe Pro Asn Trp Pro
226 605 610 615
228 ATG GCA CTC CTG ATC ACC CTC ATC GTC GTG GCG ACG CTG CCC ATC CCT 2164
229 Met Ala Leu Leu Ile Thr Leu Ile Val Val Ala Thr Leu Pro Ile Pro
230 620 625 630
232 GTG GTG TTC GTC CTG CGG CAC TTC CAC CTG CTC TCT GAT GGC TCC AAC 2212
233 Val Val Phe Val Leu Arg His Phe His Leu Leu Ser Asp Gly Ser Asn
234 635 640 645
236 ACC CTC TCC GTG TCC TAC AAG AAG GCC CGC ATG ATG AAG GAC ATC TCC 2260
237 Thr Leu Ser Val Ser Tyr Lys Lys Ala Arg Met Met Lys Asp Ile Ser
238 650 655 660
240 AAC CTG GAG GAG AAC GAT GAG ACC CGC TTC ATC CTC AGC AAG GTG CCC 2308
241 Asn Leu Glu Glu Asn Asp Glu Thr Arg Phe Ile Leu Ser Lys Val Pro
242 670 675 680
244 AGT GAG GCA CCT TCC CCC ATG CCC ACT CAC CGT TCC TAT CTG GGG CCC 2356
245 Ser Glu Ala Pro Ser Pro Met Pro Thr His Arg Ser Tyr Leu Gly Pro
246 685 690 695
248 GGC AGC ACA TCA CCC CTG GAG ACC AGC TGG AAC CCC AAT GGA CCC TAT 2404
249 Gly Ser Thr Ser Pro Leu Glu Thr Ser Trp Asn Pro Asn Gly Pro Tyr
250 700 705 710
252 GGG CGC GGC TAC CTG CTG GCC AGC ACC CCT GAG TCT GAG CTG 2446
253 Gly Arg Gly Tyr Leu Leu Ala Ser Thr Pro Glu Ser Glu Leu
254 715 720 725
256 TGACCACTGC CCAAGCCCAT GCCCGCTCTC CCCCCACCG 2485
259 (2) INFORMATION FOR SEQ ID NO: 2:
261 (i) SEQUENCE CHARACTERISTICS:
262 (A) LENGTH: 727 amino acids
263 (B) TYPE: amino acid

```

RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/923,444A

DATE: 02/06/2002

TIME: 08:56:41

Input Set : N:\Crf3\RULE60\09923444A.txt

Output Set: N:\CRF3\02062002\I923444A.raw

```

264      (D) TOPOLOGY: linear
266      (ii) MOLECULE TYPE: protein
268      (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2:
270 Met Pro Lys Asn Ser Lys Val Thr Gln Arg Glu His Ser Ser Glu His
271      1      5      10      15
273 Val Thr Glu Ser Val Ala Asp Leu Leu Ala Leu Glu Glu Pro Val Asp
274      20      25      30
276 Tyr Lys Gln Ser Val Leu Asn Val Ala Gly Glu Ala Gly Gly Lys Gln
277      35      40      45
279 Lys Ala Val Glu Glu Glu Leu Asp Ala Glu Asp Arg Pro Ala Trp Asn
280      50      55      60
282 Ser Lys Leu Gln Tyr Ile Leu Ala Gln Ile Gly Phe Ser Val Gly Leu
283      65      70      75      80
285 Gly Asn Ile Trp Arg Phe Pro Tyr Leu Cys Gln Lys Asn Gly Gly Gly
286      85      90      95
288 Ala Tyr Leu Val Pro Tyr Leu Val Leu Leu Ile Ile Ile Gly Ile Pro
289      100      105      110
291 Leu Phe Phe Leu Glu Leu Ala Val Gly Gln Arg Ile Arg Arg Gly Ser
292      115      120      125
294 Ile Gly Val Trp His Tyr Ile Cys Pro Arg Leu Gly Gly Ile Gly Phe
295      130      135      140
297 Ser Ser Cys Ile Val Cys Leu Phe Val Gly Leu Tyr Tyr Asn Val Ile
298      145      150      155      160
300 Ile Gly Trp Ser Ile Phe Tyr Phe Phe Lys Ser Phe Gln Tyr Pro Leu
301      165      170      175
303 Pro Trp Ser Glu Cys Pro Val Val Arg Asn Gly Ser Val Ala Val Val
304      180      185      190
306 Glu Ala Glu Cys Glu Lys Ser Ser Ala Thr Thr Tyr Phe Trp Tyr Arg
307      195      200      205
309 Glu Ala Leu Asp Ile Ser Asp Ser Ile Ser Glu Ser Gly Gly Leu Asn
310      210      215      220
312 Trp Lys Met Thr Leu Cys Leu Leu Val Val Trp Ser Ile Gly Gly Met
313      225      230      235      240
315 Ala Val Gly Lys Gly Ile Gln Ser Ser Gly Lys Val Met Tyr Phe Ser
316      245      250      255
318 Ser Leu Phe Pro Tyr Val Val Leu Ala Cys Phe Leu Val Arg Gly Leu
319      260      265      270
321 Leu Leu Arg Gly Ala Val Asp Gly Ile Leu His Met Phe Thr Pro Lys
322      275      280      285
324 Leu Val Lys Met Leu Asp Pro Gln Val Trp Arg Glu Val Ala Thr Gln
325      290      295      300
327 Val Phe Phe Gly Leu Gly Leu Gly Phe Gly Gly Val Ile Val Phe Ser
328      305      310      315      320
330 Ser Tyr Asn Lys Gln Asp Asn Asn Cys His Phe Asp Gly Ala Leu Val
331      325      330      335
333 Ser Phe Ile Asn Phe Phe Thr Ser Val Leu Ala Thr Leu Val Val Phe
334      340      345      350
336 Val Val Leu Gly Phe Lys Ala Asn Ile Met Asn Glu Lys Cys Val Val
337      355      360      365

```

VERIFICATION SUMMARY

PATENT APPLICATION: US/09/923,444A

DATE: 02/06/2002

TIME: 08:56:42

Input Set : N:\Crf3\RULE60\09923444A.txt

Output Set: N:\CRF3\02062002\I923444A.raw

L:28 M:220 C: Keyword misspelled or invalid format, [(A) APPLICATION NUMBER:]
L:29 M:220 C: Keyword misspelled or invalid format, [(B) FILING DATE:]